DOCUMENT RESUME

ED 050 031 SP 004 899

AUTHOR Belland, John C.; And Others
TITLE Analyzing Teacher Questions: A Comparative Evaluation of Two Observation Systems.
PUB DATE 71
NOTE 14p.; Paper presented at AERA annual meeting, New York, 1971

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS *Classroom Observation Techniques, *Comparative Analysis, *Interaction Process Analysis, *Questioning Techniques, *Reliability, Teacher Behavior, Validity
IDENTIFIERS Hough Duncan Observation System, Price Belland Observation System

ABSTRACT

General systems for analyzing instructional interaction have found the most common teacher behavior to be asking questions. This evaluation compares and contrasts two systems for analyzing teacher questions: Price-Belland, developed by the authors from the Bloom-Saunders tradition, and Hough-Duncan, modified for detailed question analysis. Comparisons were made 1) on the nature of the decision-making process required in coding, 2) on the kind of information derived from each system, and 3) on the interpretability of data displays derived from each system. In comparing data obtained from using the two observation systems simultaneously, significant differences were found between systems in the percentage of activity assigned to five of seven common categories. It is recommended that work be continued on developing a more reliable questioning-behavior analysis system. (RT)

ED050031

C75

ANALYZING TEACHER QUESTIONS: A COMPARATIVE EVALUATION OF TWO OBSERVATION SYSTEMS

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

A paper presented at

the 1971 meeting of the

American Educational Research Association

New York, New York

by

John C. Belland

Ann Belland

The Ohio State University

and

Thomas J. Price

Humboldt State University

SP004879

ANALYZING TEACHER QUESTIONS: A COMPARATIVE EVALUATION OF TWO OBSERVATION SYSTEMS

In studying the spate of instructional observation systems developed in recent years, it has become apparent that teachers spend a considerable proportion of their time asking questions. Several schemes have been proposed for analyzing questioning behavior in detail.

Gall (1970) has recently published a comprehensive review of the various methodological traditions and findings in the area of questions asked by either teachers or students. He suggested that continuing effort should be expended in developing systems which will not only describe the kinds of questions which teachers ask, but also the kinds of questions which will lead students to the achievement of various educational objectives. Beyond the examining of individual questions, Gall suggested that sequences of questions or questioning strategies might assist that investigation. He also suggested that developing analytic instruments designed for specific curricular areas would tend to make analyses of the data more productive.

Very little has been written on the nature of the decision-making process used in categorizing questioning behaviors (or any behaviors for that matter). In perhaps the bulk of observation systems, the decision is made in a fashion like the responding to a matching-test item. One has a list of categories and a series of behaviors to match with them. Sometimes the list of categories is arranged along a continuum or is arranged in a hierarchical taxonomy. When this occurs, the observer is encouraged to deal with behaviors falling between two categories with an arbitrary ground rule. Sanders (1966) and Clegg et al. (1967) have proposed observation schemes based on the Bloom Taxonomy (1956). The investigators have proposed an extension of this tradition (see Figure 1).

A few observation systems have relied on the observer making a series of very simple decisions with respect to each categorization. In such systems a final category is recorded by a series of symbols representing the choice at each decision point. These decisions are commonly binary in nature although 3- and 4-way decisions also occur and are most like the decisions made by digital computers. Taba (Simon & Boyer, 1970) and Hough-Duncan (1970) (see Figure 2) have developed systems based on such decision-making trees.

It would seem likely that the nature of the decision-making process would influence the kind of description produced by an observation system. In order to make a preliminary investigation of this idea, it was decided to categorize a set of audio recordings using the Price-Belland system and the extension of the Hough-Duncan OSIA system (see Figure 2a). These two systems have several final category elements in common and could be adjusted to the same measuring unit, each new behavior.

Figure 1

CATEGORIES OF QUESTIONS AND RESPONSES
PRICE-BELLAND QUESTION ANALYSIS SYSTEM

Code		Category
Teacher	Student	
X	SX	OTHER BEHAVIOR: irrelevant, incomplete, reinforcement, criticism, discourse
1	S1	PROCEDURAL: management, encouraging or acknowledging person to speak
2	S2	INFORMATION: recognition or recall of factual information; opinion
3	S3	TRANSLATION: change of information to another communication
4	S4	INTERPRETATION: relationships
5	S5	APPLICATION: general rule to specific case; solve a problem
6	S6	ANALYSIS: divide in parts
7	S7	SYNTHESIS: put parts together to produce unique communication
8	S8	EVALUATION: judgment; theory
9	S9	AFFECTIVE: feeling, justification, belief
+	S+	CLARIFICATION: repeat, rephrase
P		PROBING: follows incorrect answer, no answer, or incomplete answer
	-	INDEFINITE: don't know, maybe, not sure
	0	SILENCE: no answer
	Z	CONFUSION: interference

Figure 2

The Observational System for Instructional Analysis

	Teacher Behaviors		Student Behaviors	
Substantive	T1	Substantive clarification	S1	Substantive
	T2	Responds to substantive solicitation	S2	
	T3	Initiates substantive information	S3	
	T4	Solicits substantive response	S4	
Appraisal	T5	Corrective feedback	S5	Appraisal
	T6	Confirmation	S6	
	T7	Acceptance	S7	
	T8	Positive personal judgment	S8	
	T9	Negative personal judgment	S9	
Managerial	T10	Managerial clarification	S10	Managerial
	T11	Responds to managerial solicitation	S11	
	T12	Initiates managerial information	S12	
	T13	Solicits managerial response	S13	
Silence	T14	Silent covert activity	S14	Silence
	T15	Silent overt activity	S15	

Teacher or Student Behavior

X 16 Instructionally non-functional behavior

Y 17 Interaction separation designation

Categories 1-4, and 10-13 may be further categorized as a. closed or b. open.

Figure 2a

DECISIONS IN HOUGH-DUNCAN OBSERVATION SYSTEM FOR INSTRUCTIONAL ANALYSIS MODIFIED FOR DETAILED ANALYSIS OF QUESTIONING BEHAVIOR

(T) Teacher	(1,4) Solicitation	(1) Clarification	(0+) Substantive	(c) Cognitive	(1) Information
(S) Student	(x) Other	(4) Solicitation	(10+) Managerial	(a) Affective	(2) Convergent Application
				(p) Psychomotor	(3) Divergent Application
					(4) Judgment

There were many other dimensions along which analyses could be made. It was determined to informally check whether training intern-teachers in questioning techniques as well as analysis would elicit changes in both intern-teacher and student behaviors. Specifically: (a) could memory questions and responses be reduced, and (b) could intern-teachers ask questions to which students could respond appropriately?

In spite of considerable effort expended by the developers of the pre-1970 systems for analysis of teacher questions to disseminate their techniques, there has been minimal use of those systems by teachers and supervisors. If questioning is an important behavior and if analysis of questioning will lead to a more controlled and flexible use of that behavior, it is important to develop a system which will make a detailed analysis of questioning behavior possible while at the same time be sufficiently simple that it can be used by teachers and supervisors.

Objectives of the Study

In order to develop a usable but detailed scheme for analyzing classroom questions, it was decided to develop an observational system which was based on the Bloom-Sanders tradition. Use of this system was compared with use of Hough-Duncan in order to evaluate: (a) differences in qualities of data, (b) differences in decision-making while categorizing.

Methods and Techniques

The steps in this investigation were to:

1. Develop a system for coding and analyzing questioning behavior as noted above
2. Carry out a pilot study
3. Code and analyze these same data using the Hough-Duncan system
4. Develop displays of the data
5. Evaluate the data

Pilot Study. The methodology of the pilot study involved the quasi experimental one-group pre- and posttest design. The intern-teacher used different randomly-selected groups of students for practice so that the students received a minimum of practice in responding to the cues of the intern-teacher.

First, each intern-teacher randomly assigned the students in his/her class into five equal-sized groups. For the purpose of gathering pretest data, each of the ten intern-teachers conducted a questioning session with

one randomly selected group of students (G_1) for a total of 15 minutes of audio recording. An observer categorized the interaction using the Price-Belland observation system.

Each teacher then met with the observer to review the training materials in this project to facilitate the self-instructional use of these materials.

Each intern-teacher learned the observation system for recording questioning interaction. Each intern-teacher read selections from a selected bibliography when necessary for further clarification. The observation system for recording questioning interaction was considered mastered by the intern-teacher when the practice tape was categorized, without stopping, with four or less responses varying from the categorization assigned by the system developers.

Teachers arranged 4, 20-minute interview situations with groups of students not used for the pretest (G_2, G_3, G_4, G_5) in which the questioning session was audio-taped. The objective of this developmental work with different students was two-fold:

a. The intern-teacher was to ask higher-level questions.

b. The teacher was to elicit higher-level responses from the students.

Each teacher recorded the interaction on audiotape, only categorizing the last 15 minutes of the teacher and student interaction. This produced a total of 1 hour 20 minutes of development work recorded on audiotape, but just 1 hour of all interaction was categorized and diagnosed by the intern-teachers. The time allowed for this developmental work was a two week period.

For a posttest session, each teacher recorded a questioning session for 15 minutes on audiotape with the same student group used in the pretest (G_1). The observer categorized this interaction.

Observer controls. The observer who categorized the tapes in the Price-Belland system was one of the designers of that system. Reliability was maintained by conferring with the other designers before each categorization session and by recategorizing the practice tape. The criterion was the complete agreement with the previously agreed upon categories.

OSIA Coding. The observer who categorized the tapes in the OSIA extension was a student of Hough at The Ohio State University. The principal investigator worked with this observer developing the definitions of the extended categories and setting out the ground rules. All tapes were analyzed by this observer in a two-day period. No quantitative measures of reliability were made. The principal investigator has worked with OSIA since 1968. If his explanations of the categories were uniform to each observer he provided the control necessary to infer that the differences resulted from the decision-making process.

Data displays. The Price-Belland data were displayed in the form of a percentage profile (see Figure 3). No attempt was made to preserve sequence information in the display. Since the objectives of the training were that the intern-teacher should be able to reduce the number of memory questions and to elicit responses appropriate to the level of the questions, this display seemed to be sufficient.

The data displays for the extended OSIA data were frequency and percentage distributions and charting the flow of the questioning (see Figure 4). Since these materials were not used to influence the intern-teachers, it was deemed unnecessary to develop any of the other displays designed for OSIA.

Data Sources

The data were collected as tape-recorded samples of questioning behavior from ten intern-teachers in May 1970. These teachers were nearing the end of a year-long internship as part of the Elementary Master of Arts in Teaching Program at the University of North Carolina, Chapel Hill.

Findings

Price-Belland observations of the pilot study. Since there were two objectives of the training for intern-teachers in the pilot study, analysis of the data follows in that format.

> H_1: If intern-teachers are trained to recognize memory questions and are placed in practice situations designed to decrease the use of such questions, they will be able to conduct a question-asking session with a lower percentage of memory questions than they did before the training and practice.

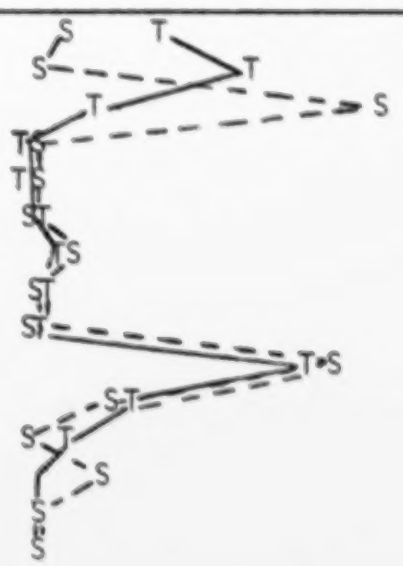
In order to test the null-hypothesis formed from H_1, the Wilcoxon Matched-Pairs Signed-Ranks test was used. When the pre and post percentages of memory questions were compared, the Wilcoxon T value was 7. This value is significant $p<.025$ when $N=10$ in a one-tailed test. Since the median percent of memory questions on the pretest was 9 and the median percent of memory questions on the posttest was 7 1/4, the null hypothesis was rejected and H_1 confirmed.

> H_2: If intern-teachers are trained to analyze questions and responses, they will be able to ask questions which will elicit appropriate category responses.

This hypothesis was not tested in a formal manner. In Table 1 the direction of the shifts in both teacher and student behaviors in each category is listed. Out of 15 categories in which change could occur, the median number of categories in which the intern-teacher and the students either shifted in the same direction or mutually remained constant was 8. The range was from

Figure 3

QUESTION AND RESPONSE PROFILE

Tabulation					Per Cent Profile
Category	Teacher		Student		0 10 20 30 40 50 60 70 80 90
	#	%	#	%	
X	13	14	3	4	
1	22	24	1	1	
2	8	9	30	38	
3	0	0	0	0	
4	0	0	0	0	
5	1	1	0	0	
6	4	4	3	4	
7	2	2	1	1	
8	1	1	0	0	
9	28	30	27	33	
+	10	11	7	9	
P	4	4	0	0	
-	-	-	6	8	
0	-	-	1	1	
Z	-	-	0	0	
Totals	93	100	79	99	

QUESTION AND RESPONSE PROFILE

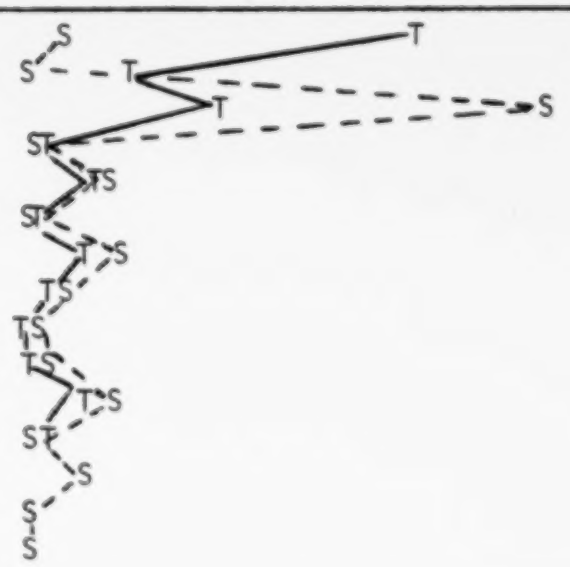
Tabulation					Per Cent Profile
Category	Teacher		Student		0 10 20 30 40 50 60 70 80 90
	#	%	#	%	
X	44	42	3	4	
1	12	11	0	0	
2	22	21	39	56	
3	2	2	1	1	
4	7	7	6	9	
5	1	1	0	0	
6	6	6	8	10	
7	3	3	2	3	
8	0	0	0	0	
9	1	1	1	1	
+	6	6	6	9	
P	2	2	0	0	
-	-	-	4	6	
0	-	-	0	0	
Z	-	-	0	0	
Totals	106		70		

Figure 4

Example of OSIA Data Display

Table 1

Changes in Questioning Behavior Before and After Training
in the Price-Belland Question Analysis System

Intern teacher number		Category X	1	2	3	4	5	6	7	8	9	+	P	-	0	Z	Common Shifts	Common Constraints	Decrease in Memory Q & R
1	T	+	-	+	+	+	0	+	+	-	-	-	-	0	0	0	7	2	Yes
	S	0	-	+	+	+	0	+	+	0	-	0	0	-	-	0			
2	T	-	-	+	0	+	-	+	0	+	+	+	-	0	0	0	7	1	Yes
	S	-	-	+	+	0	-	+	0	+	-	+	0	+	-	+			
3	T	+	-	-	0	0	+	-	-	-	-	+	-	0	0	0	7	3	Yes
	S	-	-	-	0	0	+	-	-	-	-	0	0	-	0	-			
4	T	+	-	-	0	+	+	-	+	0	+	-	-	0	0	0	5	2	Yes
			*																
	S	+	0	-	0	+	+	0	+	0	-	0	0	-	-	-			
5	T	+	+	-	0	+	0	+	+	0	-	-	+	0	0	0	6	4	No
	S	0	0	-	0	+	0	+	+	0	-	-	0	-	+	0			
6	T	+	+	-	+	+	+	+	0	-	0	-	-	0	0	0	6	1	No
	S	+	+	-	+	0	+	+	0	0	-	+	0	-	-	-			
7	T	-	-	-	0	+	+	+	+	+	0	+	+	0	0	0	6	1	Yes
			*																
	S	+	0	+	0	+	+	+	+	+	-	-	0	-	-	+			
8	T	+	-	+	0	0	-	+	0	0	-	+	-	0	0	0	5	3	Yes
			*																
	S	+	0	+	+	0	-	+	0	0	-	0	0	+	-	+			
9	T	+	0	+	0	+	+	-	-	-	+	-	-	0	0	0	7	3	No
	S	0	0	-	+	+	+	-	-	-	+	-	0	0	-	0			
10	T	0	-	+	-	+	0	+	-	0	+	+	0	0	0	0	2	3	Yes
			*																
	S	+	0	-	+	0	+	0	0	0	+	+	0	+	0	+			

+ = increase from pre to post measures
- = decrease from pre to post measures
0 = no change from pre to post measures
* = zero in both measures

5 to 10. It appeared that this was not indicative of a strong effect so H_2 could not be supported by the evidence above.

Comparisons between Price-Belland and OSIA data. In order to make the data from the two observation systems comparable, they were grouped into seven categories according to the definitions by Hough and Duncan (1970): (a) memory, (b) convergent, (c) divergent, (d) evaluation, (e) affective, and (f) clarification solicitations; and (g) other behaviors. Table 2 summarizes the data from the two systems. The hypothesis was that the percent of behavior in each category would be different in each system. The Wilcoxon Matched-Pairs Signed-Ranks Test was used to test whether there was a significant difference in the same direction between the percents in a category derived from the two systems. Using the two-tailed test, the categories memory, divergent, affective, and other each showed significant differences.

In the categories convergent and clarification, there was no lack of difference, but these differences occurred in varying directions and were ignored by the statistical test. In the evaluation category, there were so few uses of this behavior that no differences could be measured. Generally, in more than half the categories there were significant differences in percent of behavior use reported by the two systems. Thus if all the assumptions hold, it is likely that the decision-making process will influence the category outcomes.

Recommendations

Since training and practice in analysis of questions can influence question-asking behavior, it seems appropriate to continue to evolve observation systems which will yield information about the relationship of certain questions and question-asking strategies to educational objectives. In this design process, the validity of the products from various decision-making strategies will have to be studied. Either one of the two systems compared here or some other system must have a closer relationship to the reality perceived by the participants than the other systems are able to produce.

The development of a system for analyzing teacher questioning behavior and student response which provides: (a) decisive decision-making in categorizing, (b) detailed recording of the teacher-student interaction, and (c) data easily analyzed and interpreted by teacher, supervisor, and researcher would facilitate study of this most-common teaching behavior. Hopefully, this comparison of recent systems representing two traditions of questioning-behavior analysis will contribute to the development of such a system.

Table 2

Comparisons Between Price-Belland and OSIA Data

Intern Teacher #	Observation	Percent of Behavior in Teacher Question Category													
		Memory		Convergent		Divergent		Evaluation		Affective		Clarification		Other	
		PB	OSIA	PB	OSIA	PB	OSIA	PB	OSIA	PB	OSIA	PB	OSIA	PB	OSIA
1	Pre	15½	0	9	7	1½	27	½	0	0	0	9½	13	64	52½
	Post	7½	6½	22	17	2	14	0	0	½	0	4½	10½	63	51½
2	Pre	18½	0	13½	4	0	34½	0	0	0	0	2½	10	65	51
	Post	13	3	13½	9	0	21	½	0	5	6	8	8	60½	53
3	Pre	11	0	15	9½	1	29	1	1½	4	0	8½	8	60	52
	Post	10	1	14½	9½	0	25	0	0	5½	0	3½	4	66½	60½
4	Pre	4½	2	20½	23	0	10½	0	0	1	0	11	9½	62½	55
	Post	2	1½	15	28½	3½	9½	0	0	½	0	10½	11	68	51
5	Pre	6	6½	15½	6½	0	16½	0	0	15½	8½	11½	12	52	50
	Post	6½	1	19	16½	2½	17	0	1½	7½	1½	4	12	60	50
6	Pre	8½	½	17	20½	0	7½	½	0	3	0	23	13½	48½	58
	Post	10½	1	19	20	0	10	0	0	2	1	6½	10	62	58
7	Pre	8	1½	18	31	0	1½	0	0	2	0	5½	10	66½	56
	Post	7	1	17½	25	2	12	½	0	1½	0	9½	10	62½	52
8	Pre	9½	1	14½	25	0	14	0	0	8½	1	20½	7	47	52
	Post	3	1	17½	33	0	2	0	0	2½	0	11	6	65½	58
9	Pre	8	0	13	12½	3	15½	2	0	0	0	14	22	60	50
	Post	7	0	15	19	0	12	0	2	4	12	4	5	71	50
10	Pre	16½	2	12	16	½	8	½	2	½	0	10	16	60	56
	Post	11	3	16½	11½	0	9½	½	0	1½	0	10½	20½	60	55
T		1**		94.5		6.5**		20		25*		57		15.5**	

(N reduced by ties)

* p .05 in a two-tailed test
** p .01 in a two-tailed test

Note: T is the statistic from the Wilcoxon Matched-Pairs Signed-Ranks Test (Siefel, 1956, pp. 75-83).

REFERENCES

Bloom, B. S. et al. Taxonomy of Educational Objectives Handbook I. New York: David McKay, 1956.

Clegg, A. A. et al. Training Teachers to Analyze the Cognitive Level of Classroom Questions. Amherst: University of Massachusetts, 1969.

Gall, M. D. "The Use of Questions in Teaching." Review of Educational Research, Vol. 40, No. 5, December 1970.

Hough, J. B. and Duncan, J. K. Teaching: Description and Analysis. Reading, Mass.: Addison-Wesley, 1970.

Sanders, N. M. Classroom Questions--What Kinds? New York: Harper & Row, 1966.

Simon, A. & Boyer, E. G. ed. Mirrors for Behavior II. Philadelphia: Research for Better Schools, 1970.